

# Nonlinear stochastic biasing of halos: Analysis of cosmological $N$ -body simulations and perturbation theories

Masanori Sato<sup>1\*</sup> and Takahiko Matsubara<sup>1,2</sup>

<sup>1</sup> *Department of Physics, Nagoya University, Chikusa, Nagoya 464-8602, Japan and*

<sup>2</sup> *Kobayashi-Maskawa Institute for the Origin of Particles and the Universe,  
Nagoya University, Chikusa, Nagoya 464-8602, Japan*

(Dated: April 17, 2013)

It is crucial to understand and model a behavior of galaxy biasing for future ambitious galaxy redshift surveys. Using 40 large cosmological  $N$ -body simulations for a standard  $\Lambda$ CDM cosmology, we study the cross-correlation coefficient between matter and halo density field, which is an indicator of stochasticity of bias, over a wide redshift range  $0 \leq z \leq 3$ . The cross-correlation coefficient is important to extract information on the matter density field, e.g., by combining galaxy clustering and galaxy-galaxy-lensing measurements. We compare the simulation results with integrated perturbation theory (iPT) proposed by one of the present authors and standard perturbation theory (SPT) combined with a phenomenological model of local bias. The cross-correlation coefficient derived from the iPT agrees with  $N$ -body simulation results down to  $r \sim 15$  (10)  $h^{-1}$  Mpc within 0.5 (1.0) % for all redshifts and halo masses we consider. The SPT with local bias does not explain complicated behaviors on quasi-linear scales at low redshifts, while roughly reproduces general behavior of the cross-correlation coefficient on fully nonlinear scales. The iPT is powerful to predict the cross-correlation coefficient down to quasi-linear regimes with a high precision.

PACS numbers: 98.80.Es

## I. INTRODUCTION

In the standard cosmological model, known as the  $\Lambda$ CDM model, the energy density is dominated by mysterious components called as dark matter and dark energy. The correlation function of dark matter and its Fourier counterpart, the power spectrum, contain a wealth of information that can be used to determine, e.g., the dark matter, dark energy, and neutrino masses. Thus, it is very important to exploit these quantities in the large-scale structure of the universe, which is a pillar of modern observational cosmology. However, it needs to be investigated how to take account of galaxy biasing. Observable galaxies are biased relative to the underlying matter density field. The galaxy biasing is affected by nonlinear effects and is scale dependent in general. Such nonlinear effects impose a serious problem in analyzing galaxy surveys [e.g., 1–4]. Upcoming galaxy surveys such as BigBOSS<sup>1</sup> [5], Euclid [6], Subaru Prime Focus Spectrograph (PFS)<sup>2</sup> [7] and the Wide-Field Infrared Survey Telescope (WFIRST)<sup>3</sup> require the understanding of galaxy biasing with high precision and thus theoretically precise description of the galaxy biasing is a crucial issue.

Most of the direct studies of the clustering of matter on cosmological scales rely on shear-shear weak lensing, but it is also possible to extract information on the matter clustering by combining galaxy clustering and galaxy-galaxy lensing measurements [e.g., 8]. To achieve this, one has to precisely know the relation between the distribution of galaxies and the distribution of matter. An important property of the relation

is often characterized by a cross-correlation coefficient. The cross-correlation coefficient is a characteristic parameter of stochasticity [9]. Since galaxies are expected to form in dark matter halos in modern models of galaxy formation, understanding and modeling the clustering properties of the halos play an important role and are crucial first step in modeling galaxy biasing.

In this work, we examine how well known models of halo clustering reproduce the cross-correlation coefficient between matter and halo density fields obtained from  $N$ -body simulations. We consider two models of nonlinear bias: the integrated Perturbation Theory (iPT) developed by Matsubara [10] which naturally incorporates the halo bias, redshift-space distortions, nonlocal Lagrangian bias and primordial non-Gaussianity in a formalism of perturbation theory with a resummation technique based on the Lagrangian picture (see also [11, 12]), and the standard perturbation theory (SPT) combined with phenomenological model of local bias, which leads to non-trivial renormalizations of the leading order bias parameter [13]. A significant advantage of the iPT is that it is simpler and easier to calculate the power spectrum than other resummation methods even in the presence of halo bias and redshift-space distortions. The computational cost is similar to that of the SPT.

We focus not on power spectrum but on two-point correlation function, because we do not suffer from shot noise effect in the correlation function. While 2-loop corrections in the iPT generally might have an impact on the correlation function on quasi-linear scales [14], we use 1-loop iPT for simplicity in this paper.

This paper is organized as follows. We first review the theoretical predictions of power spectrum and correlation function in Section II. We describe the details of  $N$ -body simulations and a method to compute the correlation functions of matter and halos from  $N$ -body simulations in Section III. After

\*masanori@nagoya-u.jp

<sup>1</sup> <http://bigboss.lbl.gov/>

<sup>2</sup> <http://sumire.ipmu.jp/pfs/intro.html>

<sup>3</sup> <http://wfirst.gsfc.nasa.gov/>

showing the results of matter and halo correlation functions and its cross-correlation function in Section IV, we then show the main results of this paper in Section V. Finally, Section VI is devoted to our conclusion.

## II. ANALYTIC MODELS

In this section, we briefly review two theoretical models: the iPT with nonlocal bias model and the SPT with local bias model, which are compared with  $N$ -body simulation results.

### A. Predictions of integrated Perturbation Theory

We use iPT [10] to investigate how the cross-correlation coefficient behaves on quasi-linear scales for various halo masses and redshifts. It is convenient to write down the power spectrum predictions of the iPT based on multipoint propagators recently introduced in Bernardeau et al. [15]. Using the multipoint propagators  $\Gamma^{(n)}$ , one-loop power spectrum between object  $X$  and  $Y$  based on the iPT can be written as (the full derivation is given in [16])

$$P_{XY}(k) = \Pi^2(k) \left[ \hat{\Gamma}_X^{(1)}(k) \Gamma_Y^{(1)}(k) P_L(k) + \frac{k^3}{8\pi^2} \int_0^\infty dr r^2 \int_{-1}^1 dx \hat{\Gamma}_X^{(2)}(k, r, x) \hat{\Gamma}_Y^{(2)}(k, r, x) P_L(kr) P_L(k \sqrt{1+r^2-2rx}) \right], \quad (1)$$

where indices  $X$  and  $Y$  are either matter 'm' or halo 'h' in this paper,  $P_L(k)$  is the linear matter power spectrum, and vertex factor  $\Pi(k)$  is given by

$$\Pi(k) = \exp \left[ -\frac{k^2}{12\pi^2} \int dp P_L(p) \right]. \quad (2)$$

The normalized multipoint propagators  $\hat{\Gamma}^{(1)}$  and  $\hat{\Gamma}^{(2)}$  are given by

$$\hat{\Gamma}_m^{(1)}(k) = 1 + \frac{k^3}{4\pi} \int_0^\infty dr \int_{-1}^1 dx \left[ \frac{5}{21} \frac{r^2(1-x^2)^2}{1+r^2-2rx} + \frac{3}{7} \frac{(1-rx)(1-x^2)rx}{1+r^2-2rx} \right] P_L(k), \quad (3)$$

$$\hat{\Gamma}_h^{(1)}(k) = 1 + c_h^{(1)}(k) + \frac{k^3}{4\pi} \int_0^\infty dr \int_{-1}^1 dx \left\{ \frac{5}{21} \frac{r^2(1-x^2)^2}{1+r^2-2rx} + \frac{3}{7} \frac{(1-rx)(1-x^2)}{1+r^2-2rx} [rx + r^2 c_h^{(1)}(kr)] \right\} P_L(k), \quad (4)$$

$$\hat{\Gamma}_m^{(2)}(k, r, x) = \frac{x}{r} + \frac{1-rx}{1+r^2-2rx} - \frac{4}{7} \frac{1-x^2}{1+r^2-2rx}, \quad (5)$$

$$\hat{\Gamma}_h^{(2)}(k, r, x) = \frac{x}{r} \left[ 1 + c_h^{(1)}(k \sqrt{1+r^2-2rx}) \right] + \frac{1-rx}{1+r^2-2rx} \left[ 1 + c_h^{(1)}(kr) \right] - \frac{4}{7} \frac{1-x^2}{1+r^2-2rx} + c_h^{(2)}(kr, k \sqrt{1+r^2-2rx}), \quad (6)$$

where indices 'm' and 'h' denote the matter and halo, and we assume the second-order renormalized bias function  $c_h^{(2)}(\mathbf{k}_1, \mathbf{k}_2)$  depends only on the magnitudes of wavevectors,  $k_1 \equiv |\mathbf{k}_1|$  and  $k_2 \equiv |\mathbf{k}_2|$ . From Equations (3)-(6), we can easily understand that the matter result is recovered when  $c_h^{(1)} = c_h^{(2)} = 0$ . Here  $c_h^{(1)}$  and  $c_h^{(2)}$  are renormalized bias functions in Lagrangian space introduced in Matsubara [10] and obtained as

$$c_h^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n) = \frac{\int_{\nu_1}^{\nu_2} \frac{f_{MF}(\nu)}{M} \hat{c}_h^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n; \nu) d \ln \nu}{\int_{\nu_1}^{\nu_2} \frac{f_{MF}(\nu)}{M} d \ln \nu}, \quad (7)$$

for a mass range  $M_1 \leq M \leq M_2$  (see Equations 64 and 108 of Matsubara [17]), where  $\nu = \delta_c / \sigma(M)$  is a function of mass  $M$ , and  $\delta_c$  is the critical overdensity for spherical collapse. In an Einstein-de Sitter cosmology, the critical overdensity is  $\delta_c \approx 1.686$ , while it shows weak dependence on cosmology and redshifts in general cosmology [18, 19], and thus we use Henry [19] fitting formula to include cosmological depen-

dence. The function  $\sigma(M)$  is the root-mean-square linear density field smoothed with a top hat filter of radius  $R$  enclosing an average mass  $M = \rho_0 4\pi R^3 / 3$ ;

$$\sigma^2(M) = \int \frac{k^2 dk}{2\pi^2} W^2(kR) P_L(k), \quad (8)$$

with

$$W(x) = \frac{3j_1(x)}{x} = \frac{3}{x^3} (\sin x - x \cos x), \quad (9)$$

where  $\rho_0$  is the mean matter density of the universe and  $j_1(x)$  is the first-order spherical Bessel function.  $f_{MF}(\nu)$  is the scaled differential mass function defined as [20]

$$f_{MF}(\nu) = \frac{M}{\rho_0} n(M) \frac{dM}{d \ln \nu}, \quad (10)$$

where  $n(M)$  is the comoving number density of halos with mass  $M$ . The quantity  $f_{MF}(\nu)$  is frequently used in the literature and there have been several analytic predictions [21–23] and fitting formulas [e.g., 20, 24–29]. In this paper, we

use Bhattacharya et al. [29] fitting formula for mass function which shows better agreement with our simulations [30].  $\hat{c}_n^L$  is given as (see, Equations 92, 95, and 96 in Matsubara [17])

$$\hat{c}_h^{(n)}(\mathbf{k}_1, \dots, \mathbf{k}_n; \nu) = b_n^L(M) W(k_1 R) \cdots W(k_n R) + \frac{A_{n-1}(M)}{\delta_c^n} \frac{d}{d \ln \sigma(M)} [W(k_1 R) \cdots W(k_n R)], \quad (11)$$

with

$$A_0(M) = 1, \quad (12)$$

$$A_1(M) = 1 + \delta_c b_1^L(M), \quad (13)$$

where  $b_n^L$  is the Lagrangian bias function for the halo bias.

Theoretical two-point correlation function can be expressed in terms of the power spectrum as

$$\xi_{XY}(r) = \int \frac{k^2 dk}{2\pi^2} \frac{\sin(kr)}{kr} P_{XY}(k). \quad (14)$$

### B. Standard perturbation theory with local bias model

In the SPT formalism, we consider the *local deterministic nonlinear biasing* model. Following Fry and Gaztanaga [31], we restrict the consideration on large scales in Eulerian space and assume that the halo density can be described by a smoothed function  $\mathcal{F}(\delta_m)$  that depends only on the matter density. We can expand  $\mathcal{F}$  in a Taylor series around  $\delta_m$  such that

$$\delta_h = \mathcal{F}(\delta_m) = \sum_{n=1}^{\infty} \frac{b_n^E}{n!} \delta_m^n, \quad (15)$$

where  $\delta_m$  is the nonlinear matter density. We then combine this expansion with SPT, which expands the matter density perturbations into a series  $\delta_m = \delta_m^{(1)} + \delta_m^{(2)} + \cdots$ , where  $\delta_m^{(1)}$  is the linear density field and  $\delta_m^{(n)}$  is of order  $[\delta_m^{(1)}]^n$ . At the next-to-leading order, we can obtain the auto- and cross-power spectrum of halos as [13, 32]

$$P_{hh}(k) = b_1^2 P_{NL}(k) + 2b_1 b_2 A(k) + \frac{b_2^2}{2} B(k) + N, \quad (16)$$

$$P_{hm}(k) = b_1 P_{NL}(k) + b_2 A(k), \quad (17)$$

where  $b_1$  and  $b_2$  are the renormalized bias parameters,  $N$  is the renormalized shot noise, and  $P_{NL}(k)$  is the nonlinear matter power spectrum.  $b_1$  and  $b_2$  should be determined empirically or treated as free parameters. In this paper, we will examine both cases in Section V. The terms  $A(k)$  and  $B(k)$  can be obtained as

$$A(k) = \int \frac{d^3 q}{(2\pi)^3} F_2(q, \mathbf{k} - \mathbf{q}) P_L(q) P_L(|\mathbf{k} - \mathbf{q}|), \quad (18)$$

$$B(k) = \int \frac{d^3 q}{(2\pi)^3} P_L(q) [P_L(|\mathbf{k} - \mathbf{q}|) - P_L(q)], \quad (19)$$

where  $F_2$  is the 2nd-order mode coupling kernel in SPT:

$$F_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{1}{2} \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2} \left( \frac{k_1}{k_2} + \frac{k_2}{k_1} \right) + \frac{2}{7} \left( \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2} \right)^2. \quad (20)$$

Taking Fourier transforms, we then obtain corresponding correlation functions given by

$$\xi_{hh}(r) = b_1^2 \xi_{NL}(r) + 2b_1 b_2 A(r) + \frac{b_2^2}{2} B(r), \quad (21)$$

$$\xi_{hm}(r) = b_1 \xi_{NL}(r) + b_2 A(r), \quad (22)$$

where  $\xi_{NL}$  is the nonlinear matter correlation function,  $A(r)$  and  $B(r)$  are the Fourier transform of  $A(k)$  and  $B(k)$ . Note that  $B(r) = \xi_L^2(r) - \sigma_c^2 \delta_D(r)$  where  $\xi_L(r)$  is the linear matter correlation function,  $\sigma_c^2 = \int d^3 q P_L^2(q)/(2\pi)^3$ , and  $\delta_D(r)$  is the Dirac delta function.

## III. N-BODY SIMULATIONS

### A. Simulation parameters

To obtain accurate predictions of the cross-correlation coefficient, we resort to use the high-resolution  $N$ -body simulations of structure formation. To perform the  $N$ -body simulations, we use a publicly available tree-particle mesh code, *Gadget2* [33]. We adopt the standard  $\Lambda$ CDM model with the matter density  $\Omega_m = 0.265$ , the baryon density  $\Omega_b = 0.0448$ , the dark energy density  $\Omega_\Lambda = 0.735$  with equation of state parameter  $w = -1$ , the spectral index  $n_s = 0.963$ , the variance of the density perturbations in a sphere of radius  $8h^{-1}\text{Mpc}$   $\sigma_8 = 0.80$ , and the Hubble parameter  $h = 0.71$ . These cosmological parameters are consistent with the Wilkinson Microwave Anisotropy Probe 7yr results [34]. We performed two types of simulations both with  $N_p = 1024^3$  particles in cubic boxes. First type has a side  $L_{\text{box}} = 1000h^{-1}\text{Mpc}$  with softening length  $r_s$  being  $50h^{-1}\text{kpc}$ , and the second type has a side  $L_{\text{box}} = 2000h^{-1}\text{Mpc}$  with softening length  $r_s$  being  $100h^{-1}\text{kpc}$ . These two types are named as L1000 and L2000, respectively. The initial conditions are generated based on the 2nd-order Lagrangian perturbation theory [35, 36] with the initial linear power spectrum calculated by CAMB [37]. The initial redshift is set to  $z_{\text{ini}} = 36$  for L1000 and  $z_{\text{ini}} = 31$  for L2000. We perform  $N_{\text{run}} = 30$  and 10 realizations for L1000 and L2000, respectively. We summarize simulation parameters in Table I. The L1000 simulations used in this paper are the same as L1000 used in Sato and Matsubara [30].

We store outputs at  $z = 3.0, 2.0, 1.0, 0.5, 0.3$ , and 0 and identify halos for each output using a Friends-of-Friends (FOF) group finder with linking length of 0.2 times the mean separation [38]. We select halos in which the number of particles,  $N_p$ , is equal to or larger than 20 which corresponds to the halos with masses  $1.37 \times 10^{12} h^{-1} M_\odot$  for L1000 and  $1.10 \times 10^{13} h^{-1} M_\odot$  for L2000. Then we divide halos into five mass bins to keep track of their different clustering properties. The average number and mass of halos among realizations for redshifts are listed in Table II. The halo catalogs of Bin 4 in

TABLE I: Parameters in high- and low-resolution  $N$ -body simulations: the matter density  $\Omega_m$ , the dark energy density  $\Omega_\Lambda$ , the baryon density  $\Omega_b$ , the Hubble parameter  $h$ , the spectral index  $n_s$ , the variance of the density perturbations at  $8h^{-1}\text{Mpc}$   $\sigma_8$ , the box size  $L_{\text{box}}$ , the number of particles  $N_p$ , the initial redshift  $z_{\text{ini}}$ , the softening length  $r_s$ , and the number of realizations  $N_{\text{run}}$ .

Name	$\Omega_m$	$\Omega_\Lambda$	$\Omega_b$	$h$	$n_s$	$\sigma_8$	$L_{\text{box}}$	$N_p$	$z_{\text{ini}}$	$r_s$	$N_{\text{run}}$
L1000 (high resolution)	0.265	0.735	0.0448	0.71	0.963	0.80	$1000h^{-1}\text{Mpc}$	$1024^3$	36	$50h^{-1}\text{kpc}$	30
L2000 (low resolution)	0.265	0.735	0.0448	0.71	0.963	0.80	$2000h^{-1}\text{Mpc}$	$1024^3$	31	$100h^{-1}\text{kpc}$	10

TABLE II: Properties of halo catalogs of high- and low-resolution  $N$ -body simulations for each mass bin. We use the 5 halo catalogs abbreviated to “Bin 1”, ..., “Bin 5”.  $\bar{N}_h$  and  $\bar{M}_h$  are the average halo numbers and average halo masses at various redshifts.

L1000								
Bin 1 $1.37 \leq M_h/(10^{12}h^{-1}M_\odot) < 4.11$			Bin 2 $4.11 \leq M_h/(10^{12}h^{-1}M_\odot) < 12.32$			Bin 3 $1.23 \leq M_h/(10^{13}h^{-1}M_\odot) < 3.70$		
$z$	$\bar{N}_h$	$\bar{M}_h[h^{-1}M_\odot]$	$z$	$\bar{N}_h$	$\bar{M}_h[h^{-1}M_\odot]$	$z$	$\bar{N}_h$	$\bar{M}_h[h^{-1}M_\odot]$
3.0	$3.56 \times 10^5$	$2.08 \times 10^{12}$	3.0	$4.14 \times 10^4$	$6.06 \times 10^{12}$	3.0	$2.50 \times 10^3$	$1.69 \times 10^{13}$
2.0	$9.88 \times 10^5$	$2.17 \times 10^{12}$	2.0	$1.97 \times 10^5$	$6.42 \times 10^{12}$	2.0	$2.75 \times 10^4$	$1.83 \times 10^{13}$
1.0	$1.73 \times 10^6$	$2.23 \times 10^{12}$	1.0	$4.98 \times 10^5$	$6.69 \times 10^{12}$	1.0	$1.24 \times 10^5$	$1.96 \times 10^{13}$
0.5	$1.95 \times 10^6$	$2.24 \times 10^{12}$	0.5	$6.25 \times 10^5$	$6.78 \times 10^{12}$	0.5	$1.89 \times 10^5$	$2.00 \times 10^{13}$
0.3	$1.99 \times 10^6$	$2.25 \times 10^{12}$	0.3	$6.60 \times 10^5$	$6.81 \times 10^{12}$	0.3	$2.12 \times 10^5$	$2.01 \times 10^{13}$
0	$2.02 \times 10^6$	$2.25 \times 10^{12}$	0	$6.94 \times 10^5$	$6.84 \times 10^{12}$	0	$2.39 \times 10^5$	$2.03 \times 10^{13}$

L2000					
Bin 4 $1.23 \leq M_h/(10^{13}h^{-1}M_\odot) < 3.70$			Bin 5 $3.70 \leq M_h/(10^{13}h^{-1}M_\odot) < 11.09$		
$z$	$\bar{N}_h$	$\bar{M}_h[h^{-1}M_\odot]$	$z$	$\bar{N}_h$	$\bar{M}_h[h^{-1}M_\odot]$
3.0	$2.09 \times 10^4$	$1.70 \times 10^{13}$	3.0	$3.85 \times 10^3$	$4.70 \times 10^{13}$
2.0	$2.30 \times 10^5$	$1.84 \times 10^{13}$	2.0	$1.60 \times 10^4$	$5.17 \times 10^{13}$
1.0	$1.05 \times 10^6$	$1.97 \times 10^{13}$	1.0	$1.88 \times 10^5$	$5.70 \times 10^{13}$
0.5	$1.61 \times 10^6$	$2.02 \times 10^{13}$	0.5	$3.96 \times 10^5$	$5.92 \times 10^{13}$
0.3	$1.80 \times 10^6$	$2.03 \times 10^{13}$	0.3	$4.89 \times 10^5$	$5.99 \times 10^{13}$
0	$2.03 \times 10^6$	$2.04 \times 10^{13}$	0	$6.17 \times 10^5$	$6.07 \times 10^{13}$

L2000 is constructed so that the halo mass range is the same as that of Bin 3 in L1000, as shown in Table II. Since volume of L2000 simulations is bigger than that of L1000 simulations, number of halos with a certain mass are larger for L2000 simulations.

### B. Analysis: Two-point correlation functions

To calculate the two-point correlation function of dark matters from  $N$ -body simulations, we first randomly choose number of particles  $N_{p,r} = 196^3$  and  $256^3$  for L1000 and L2000. For dark matter halos, we use all halos in each bin. Then we directly count the  $N$ -body particle and/or halos to calculate two-point correlation function instead of using method with Fast Fourier Transform [30]. We choose  $r_i$  to be the center of the  $i$ -th bin, i.e.,  $r_i = (r_i^{\min} + r_i^{\max})/2$ , where  $r_i^{\min}$  and  $r_i^{\max}$  are the minimum and maximum distances of the  $i$ -th bin.

The shot noise corrections in the halo power spectrum are subtle. If the dark matter halos are regarded as a Poisson process, we can easily subtract shot noise effect using number density of halos  $\bar{n}_h$ . However Smith et al. [39] found that this standard correction method is not exactly correct for halos, particularly for those of large mass. This is probably because in order to identify halos using the FOF algorithm, we automatically impose that distances between halos are larger than the sum of their radii, or they would have been linked as bigger halos. Thus, the shot noise effect is scale dependent and it is difficult to correctly subtract the effect of shot noise. Therefore, we use the correlation function instead of using the power spectrum, because the shot noise effect in the correlation function is weaker than that in the power spectrum.

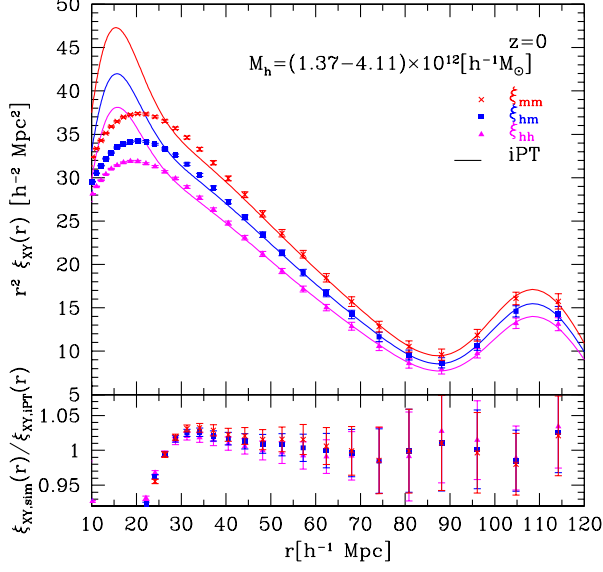


FIG. 1: *Top panel:* Correlation functions of matter, halo, and their cross-correlation function at redshift  $z = 0$ , multiplied by a factor of  $r^2$ . For halo mass range, we consider  $1.37 \leq M_h/(10^{12} h^{-1} M_\odot) < 4.11$  (Bin 1). The symbols are the results obtained from  $N$ -body simulations. The solid lines are the results of integrated perturbation theory (iPT) [10]. *Bottom panel:* Ratio of correlation functions measured from  $N$ -body simulations to those from the iPT.

#### IV. CORRELATION FUNCTIONS

Before presenting the results of cross-correlation coefficient, we compare the  $N$ -body simulation results with the iPT for correlation functions themselves.

Figure 1 shows the results for the correlation functions of matter and halos, and their cross-correlation function at  $z = 0$ . We use halo catalog of “Bin 1” shown in Table I. The amplitude of halo-halo correlation is smaller than that of matter-matter correlation, because the halo bias  $b_1^E$  in this halo range is 0.904 (less than 1). The error bars describe the  $1-\sigma$  error on the mean values obtained from 30 realizations. The error bars increase on large scales because of the finite size of the simulation box. The iPT predictions agree with  $N$ -body simulation results down to  $r \sim 25 h^{-1} \text{Mpc}$  within a few percent for all correlations. In Section V, we will see that a range of a few percent-level agreement in cross-correlation coefficient is extended more than that in the correlation functions.

#### V. CROSS-CORRELATION COEFFICIENT

In the framework of local biasing model, density field of galaxies and their halos should be a *stochastic* function of the underlying dark matter density field [9]. The stochasticity is very weak on large scales, while it becomes more important on small scales [40–42].

One of the characteristic parameters of stochasticity is the

cross-correlation coefficient between matter and halo density fields, defined as

$$r_{cc}(r) = \frac{\xi_{hm}(r)}{\sqrt{\xi_{mm}(r)\xi_{hh}(r)}}, \quad (23)$$

where  $\xi_{mm}(r)$ ,  $\xi_{hh}(r)$ , and  $\xi_{hm}(r)$  are the matter and halo auto-correlation functions, and their cross-correlation function, respectively. The cross-correlation is the measure of the statistical coherence of the two fields [43–47]. If any scale-dependent, deterministic, linear-bias model is assumed, we have  $r_{cc} = 1$ . Therefore, deviations of cross-correlation coefficient from unity would arise due to both the nonlinearity and stochasticity of bias.

Figure 2 shows the cross-correlation coefficient between matter and halo density fields at  $z = 0, 0.3, 0.5, 1.0, 2.0$  and  $3.0$ . The cross, square, and triangle symbols are the  $N$ -body simulation results measured from 30 realizations for halo masses  $1.37 \leq M_h/(10^{12} h^{-1} M_\odot) < 4.11$  (Bin 1),  $4.11 \leq M_h/(10^{12} h^{-1} M_\odot) < 12.32$  (Bin 2), and  $1.23 \leq M_h/(10^{13} h^{-1} M_\odot) < 3.70$  (Bin 3), respectively. The error bars describe the  $1-\sigma$  error on the mean value obtained from 30 realizations. We do not plot the results in which sum of the  $1-\sigma$  error bars in a range of  $5 \leq r_i/(h^{-1} \text{Mpc}) \leq 100$  is larger than 0.12, i.e.,  $\sum_{5 \leq r_i \leq 100} \sigma_i > 0.12$ . It should be noted that halos in each bin are more biased as increasing redshift, because we impose the same halo mass ranges for each bin. The solid curves show the iPT predictions. The iPT obtains good agreements with simulation results down to  $r \sim 15 h^{-1} \text{Mpc}$  within a range of error bars for all redshifts and halo mass ranges we have considered. Particularly at  $z = 3.0$ , the iPT well reproduces the simulation result down to  $r \sim 6 h^{-1} \text{Mpc}$ . Below, we will see fractional differences between simulation results and theoretical predictions in Figure 5, to discuss the percentage error. The difference between the iPT and simulation results on small scales probably comes from the fact that the iPT breaks down on small scales (see, Figure 1) [30, 48]. One can see that the iPT prediction on small scales are almost flat unlike the simulation results. This is probably because the asymptotic behaviors of the correlation functions based on the iPT are almost the same each other (see, Figure 1), and the iPT should not be applied on such small scales at any rate.

We also plot a simple model derived from Equations (21) and (22) as dotted curves and it is expressed as [32]

$$r_{cc}(r) = 1 - \frac{1}{4} \left( \frac{b_2}{b_1} \right)^2 \xi_L(r), \quad (24)$$

by using following approximations:  $A(r) \ll \xi_{NL}(r)$  and  $B(r) \ll \xi_{NL}(r)$ .

To empirically estimate  $b_1$  and  $b_2$ , we use general relations between local bias parameters in Lagrangian space and Eulerian space, which are derived in spherical collapse model as [10]

$$b_1^E = b_1^L + 1, \quad (25)$$

$$b_2^E = \frac{8}{21} b_1^L + b_2^L, \quad (26)$$



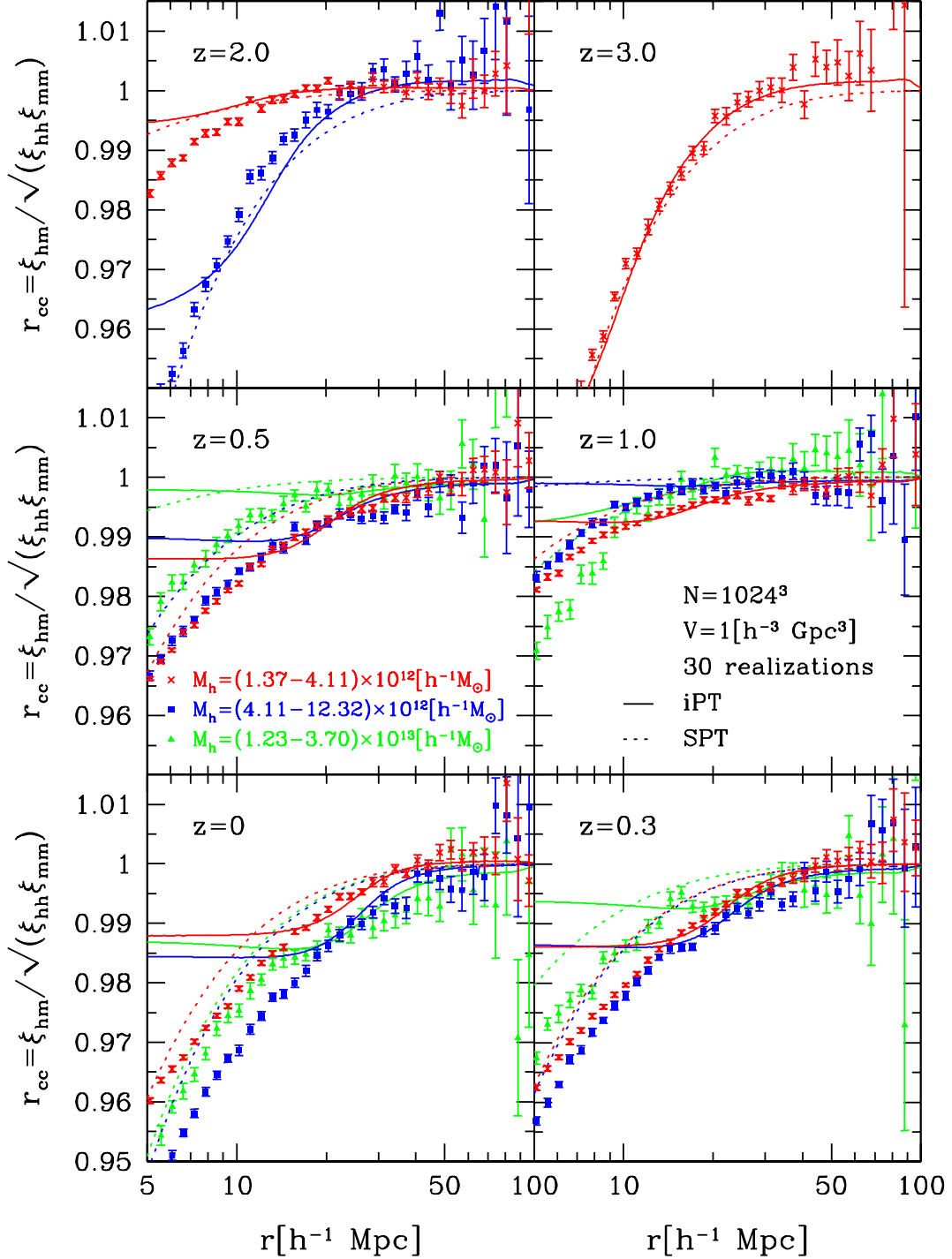


FIG. 2: The cross-correlation coefficient between matter and halo density fields at  $z = 0, 0.3, 0.5, 1.0, 2.0$ , and  $3.0$ . For halo mass ranges, we consider  $1.37 \leq M_h/(10^{12} h^{-1} M_\odot) < 4.11$  (Bin 1),  $4.11 \leq M_h/(10^{12} h^{-1} M_\odot) < 12.32$  (Bin 2), and  $1.23 \leq M_h/(10^{13} h^{-1} M_\odot) < 3.70$  (Bin 3). We do not plot the results in which sum of the  $1\text{-}\sigma$  error bars in a range of  $5 \leq r_i/(h^{-1} \text{Mpc}) \leq 100$  is larger than 0.12, i.e.,  $\sum_{5 \leq r_i \leq 100} \sigma_i > 0.12$ . The symbols are the results measured from  $N$ -body simulations. The solid lines are the results of integrated perturbation theory (iPT) [10] while the dotted lines are the results of simple model derived from standard perturbation theory with local bias model (Equation 24). To empirically estimate  $b_1$  and  $b_2$ , we use the relation in Equations (25) and (26) and then simply substitute  $b_1^E$  and  $b_2^E$  with  $b_1$  and  $b_2$ .

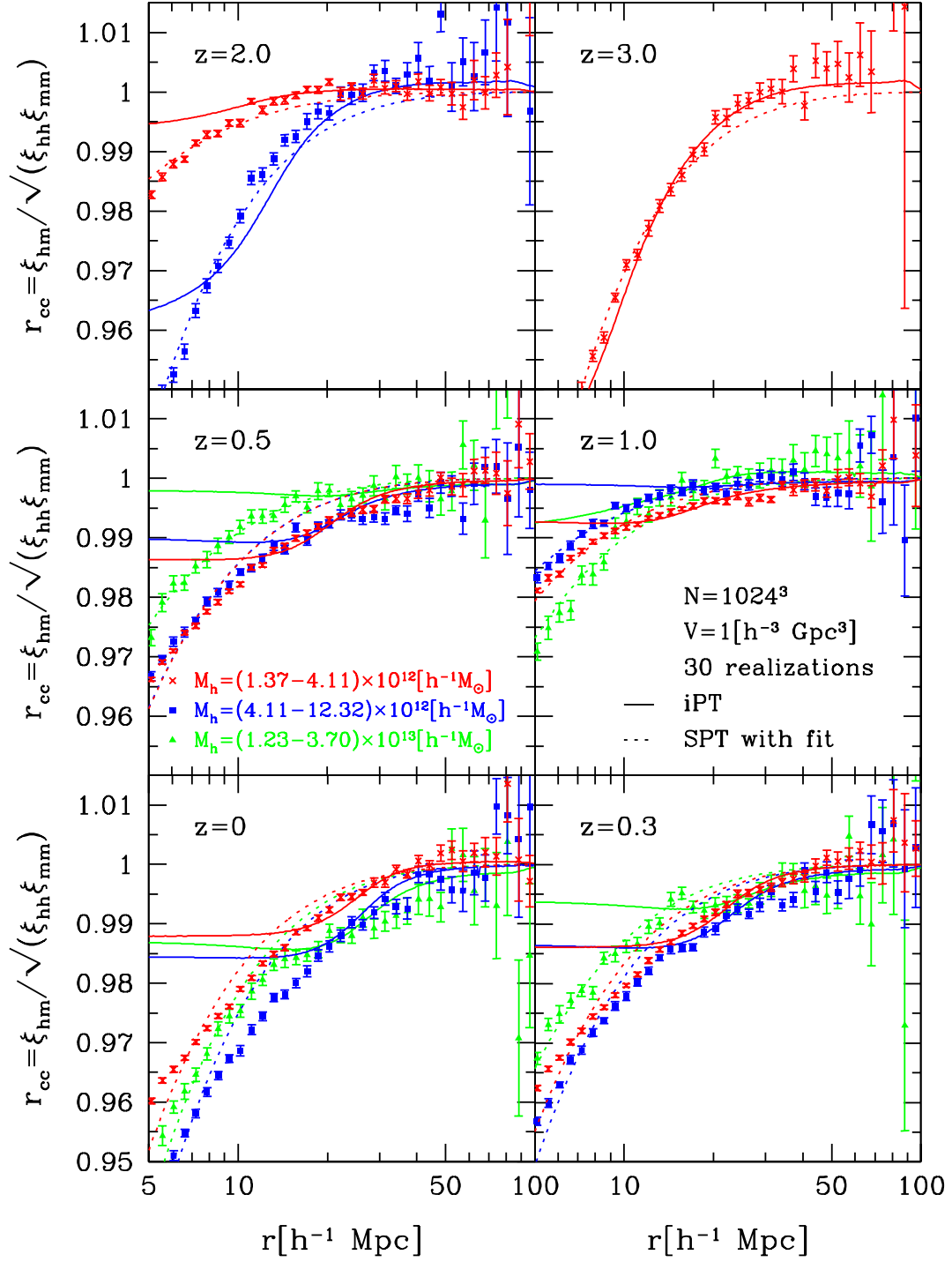


FIG. 3: Same as Figure 2, but for the dotted lines, we fit  $b_2/b_1$  to the  $N$ -body simulation results using chi-square fit.

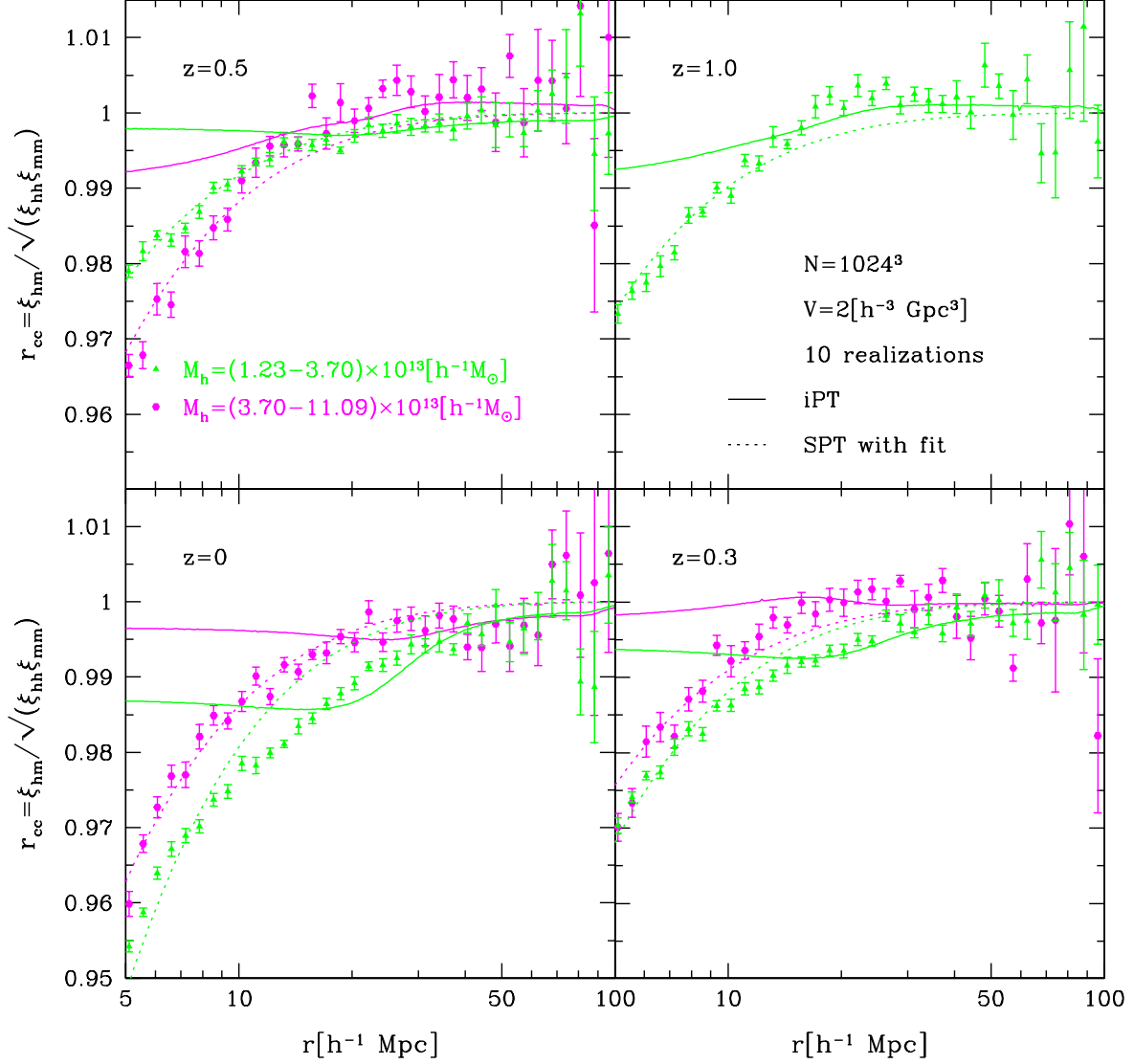


FIG. 4: The cross-correlation coefficients with halo mass ranges:  $1.23 \leq M_h/(10^{13} h^{-1} M_\odot) < 3.70$  (Bin 4), and  $3.70 \leq M_h/(10^{13} h^{-1} M_\odot) < 11.09$  (Bin 5), given at redshift  $z = 0, 0.3, 0.5$ , and  $1.0$ . We do not show the results in which sum of the  $1\text{-}\sigma$  error bars in a range of  $5 \leq r_i/(h^{-1}\text{Mpc}) \leq 100$  is larger than  $0.12$ . The triangle and circle symbols are the results of  $N$ -body simulations. The solid lines correspond to results of integrated perturbation theory (iPT) [10] while the dotted lines correspond to results of standard perturbation theory with fitted bias model (Equation 24).

where  $b_1^E$  and  $b_2^E$  are Eulerian bias parameters. Note that both the Eulerian bias parameters  $b_n^E$  and the Lagrangian bias parameters  $b_n^L$  are local and independent on scales. In this phenomenological model, we simply substitute  $b_1^E$  and  $b_2^E$  with  $b_1$  and  $b_2$ . To calculate  $b_1^L$  and  $b_2^L$ , we use [12]

$$b_n^L = \frac{(-1)^n}{\delta_c^n} \frac{\int_{M_1}^{M_2} v^n \frac{d^n f_{\text{MF}}(v)}{dv^n} \frac{d \ln \sigma(M)}{dM} \frac{dM}{M}}{\int_{M_1}^{M_2} f_{\text{MF}}(v) \frac{d \ln \sigma(M)}{dM} \frac{dM}{M}}. \quad (27)$$

for halos in a mass range  $M_1 \leq M \leq M_2$ . The simple model (Equation 24) with the above estimates of bias parameters shows better agreement with simulations for higher redshifts (i.e., more biased halos). We find that the cross-correlation coefficients of halos with  $b_1^L \gtrsim 2$  are well described in this method over all scales we considered. For lower redshifts (i.e., less biased halos), the simple model deviates more from the simulation results.

Meanwhile, when  $b_1$  and  $b_2$  are treated as free parameters, we fit  $b_2/b_1$  to the simulation results using chi-square fit. The result is shown as dotted lines in Figure 3. Other lines and



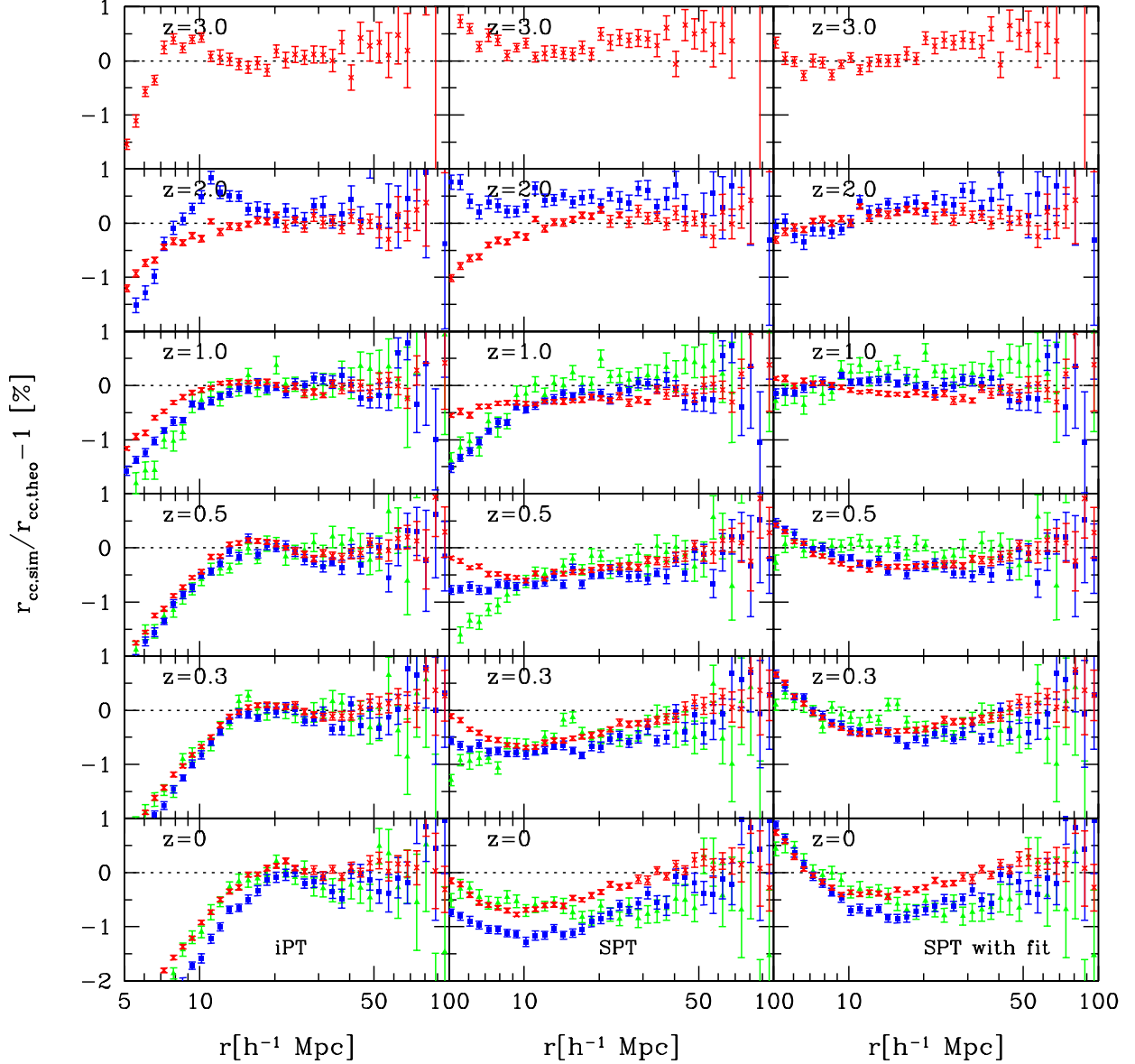


FIG. 5: Fractional differences between  $N$ -body results and theoretical predictions are shown in percents, at redshifts  $z = 0, 0.3, 0.5, 1.0, 2.0$ , and  $3.0$ , from *bottom to top*. *Left panels*: iPT, *middle panels*: SPT, and *right panels*: SPT with fitting. The red, blue, and green symbols are the results of Bin 1, 2, and 3.

symbols are the same as the Figure 2. Fittings are done in a range of  $5 \leq r/(h^{-1}\text{Mpc}) \leq 70$ . In the fitting case, an improvement from the above empirical method is little for cases of high bias, but is important for cases of low bias. The simple model with fitted bias replicates the simulation results over all scales at  $1 \leq z \leq 3$ . We can see that the cross-correlation coefficients estimated from  $N$ -body simulations have complicated behaviors in quasi-linear regimes at low redshifts, which cannot be described in the simple model. We will describe the percentage error in Figure 5 later.

Figure 4 shows the results for the cross-correlation coefficient of large halos with mass ranges:  $1.23 \leq M_h/(10^{13}h^{-1}M_\odot) < 3.70$  (Bin 4) and  $3.70 \leq M_h/(10^{13}h^{-1}M_\odot) < 11.09$  (Bin 5) at redshift  $z = 0, 0.3, 0.5$ , and  $1.0$ . The triangle and circle symbols are the simulation results of Bin 4 and Bin 5 estimated from 10 realizations of L2000. The solid and dotted lines are the predictions of the iPT and SPT with fitted bias, respectively. As in Figures 2 and 3, the iPT shows nice agreement with the simulation results on large scales even in large halo masses.

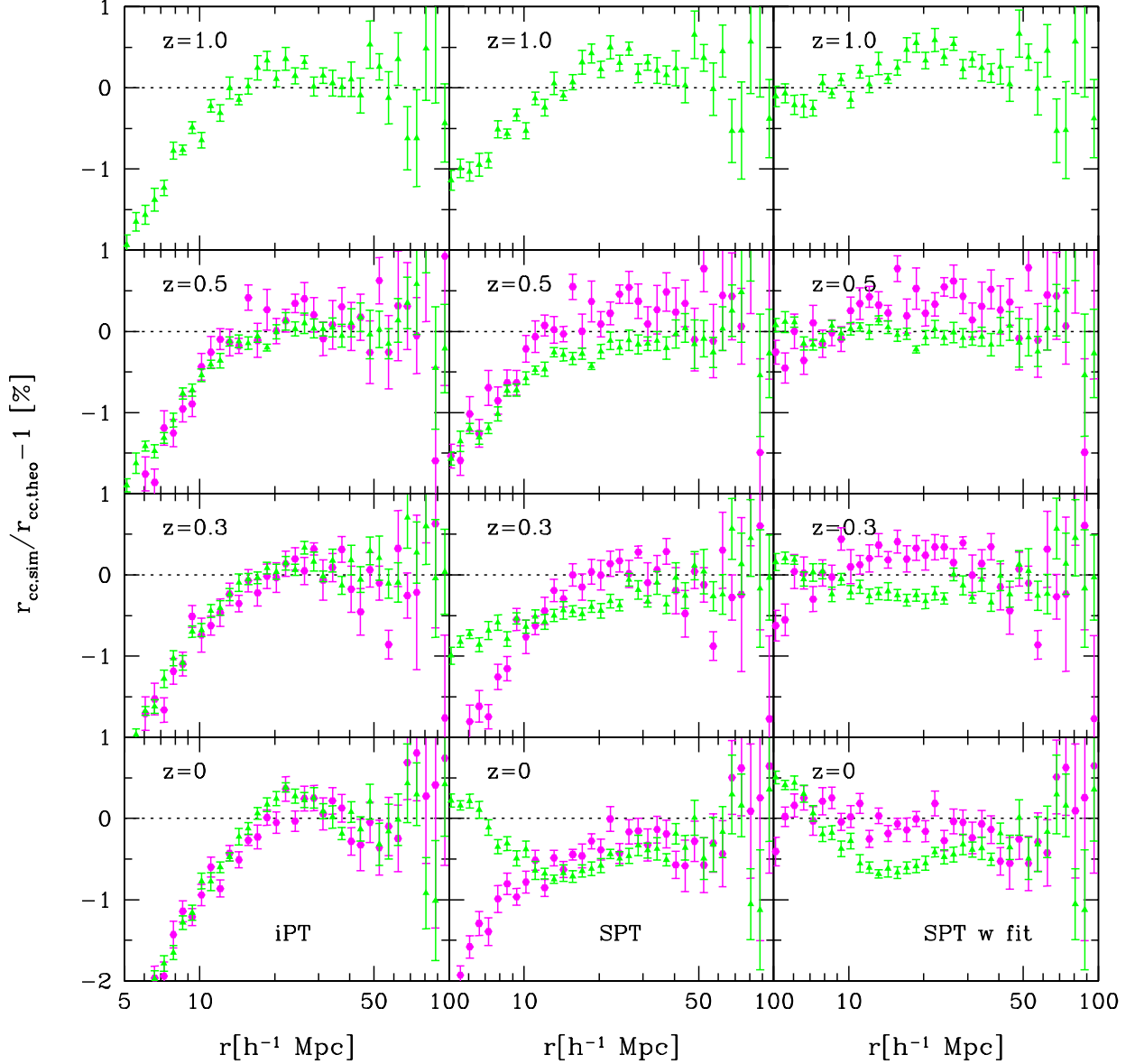


FIG. 6: Same as Figure 5, but the results of Bin 4 (green triangle) and 5 (magenta circle) at  $z = 0, 0.3, 0.5$  and  $1.0$ .

The simple model with fitting also reproduces the simulation results for large halo masses. However, fitting values of  $b_1/b_2$  are not, in general, the same as those obtained from other statistics such as power spectrum and bispectrum, because  $b_1$  and  $b_2$  are renormalized.

To clarify how well theoretical models predict the  $N$ -body results, we plot fractional differences between  $N$ -body simulation results and theoretical predictions,  $[r_{\text{cc},\text{sim}}(r) - r_{\text{cc},\text{theo}}(r)]/r_{\text{cc},\text{theo}}(r)$ , as shown in Figures 5 and 6. These Figures show that the iPT agrees with simulation results down to  $r \sim 15$  ( $10$ ) $h^{-1}\text{Mpc}$  within  $0.5$  ( $1.0$ ) % for all redshifts and halo masses we considered. It should be noted that the iPT

does not have any fitting parameter. The SPT with empirically determined bias reproduces  $N$ -body simulation results down to  $r \sim 10h^{-1}\text{Mpc}$  within a percent-level for all redshifts except for  $z = 0$  (see, Figure 5). In the SPT with bias determined by fitting, a percent-level agreement is achieved over wide separation angles for all redshifts. However, the fitted parameters  $b_1$  and  $b_2$  are different from  $b_1^E$  and  $b_2^E$  which can be determined by other method. e.g., power spectrum and bispectrum.

## VI. CONCLUSION

In this paper, we have used 40 large cosmological  $N$ -body simulations of the standard  $\Lambda$ CDM cosmology to investigate the cross-correlation coefficient between halo and matter density fields over a wide redshift range. The cross-correlation coefficient is crucial to extract information of the matter density field by combining galaxy clustering and galaxy-galaxy-lensing measurements. Since the first attempt to detect galaxy-galaxy lensing [49], its ability to constrain cosmological parameters has been shown [8].

We compare the simulation results with theoretical predictions of the iPT and simple models of bias with SPT. The iPT predicts the simulation results down to  $r \sim 15 (10)h^{-1}\text{Mpc}$  within 0.5 (1.0) % for all redshifts and halo masses we considered. To improve the prediction, the 2-loop correction to the iPT might be important. In the SPT with local bias model, bias parameters are renormalized and therefore they are determined empirically or treated as free parameters. The SPT with empirically determined biases with spherical collapse model shows better agreement with simulations for more biased halos on small scales, although this model does not reproduce complicated behaviors of the simulation results on quasi-linear scales at low redshifts. The SPT with biases determined by fitting improves the predictions but the situation is almost the same at low redshift. Thus, the iPT accurately predicts the cross-correlation coefficient as long as quasi-linear scales are considered.

Let us finally comment on convolution Lagrangian perturbation theory (CLPT), which is recently proposed by Carlson

et al. [48]. The CLPT applies additional resummations on top of the simple LRT (restricted iPT with local Lagrangian bias), and its prediction significantly improves the simple LRT for correlation function in real and redshift spaces on small scales. Therefore, it might be possible that the CLPT gives a better prediction for the cross-correlation coefficient between mass and halos and agrees with simulation results on small scales. Although it is important to examine how well the CLPT predicts, we leave it for future work.

In this paper we focus on fundamental features of bias stochasticity by the methods of numerical simulations and theoretical models. We believe the results of this paper could be a crucial first step to understand the galaxy biasing for future precision cosmology.

## Acknowledgments

We thank Uroš Seljak for useful comments. M.S. is supported by a Grant-in-Aid for the Japan Society for Promotion of Science (JSPS) fellows. T.M. acknowledges support from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT), Grant-in-Aid for Scientific Research (C), No. 24540267, 2012. This work is supported in part by a Grant-in-Aid for Nagoya University Global COE Program, “Quest for Fundamental Principles in the Universe: from Particles to the Solar System and the Cosmos”, from the MEXT of Japan. Numerical computations were in part carried out on COSMOS provided by Kobayashi-Maskawa Institute for the Origin of Particles and the Universe, Nagoya University.

- 
- [1] M. R. Blanton, D. Eisenstein, D. W. Hogg, and I. Zehavi, *ApJ* **645**, 977 (2006).
  - [2] W. J. Percival, R. C. Nichol, D. J. Eisenstein, J. A. Frieman, M. Fukugita, J. Loveday, A. C. Pope, D. P. Schneider, A. S. Szalay, M. Tegmark, et al., *ApJ* **657**, 645 (2007).
  - [3] A. G. Sánchez and S. Cole, *MNRAS* **385**, 830 (2008).
  - [4] C. Blake, S. Brough, M. Colless, C. Contreras, W. Couch, S. Croom, T. Davis, M. J. Drinkwater, K. Forster, D. Gilbank, et al., *MNRAS* **415**, 2876 (2011).
  - [5] D. Schlegel, F. Abdalla, T. Abraham, C. Ahn, C. Allende Prieto, J. Annis, E. Aubourg, M. Azzaro, S. B. C. Baltay, C. Baugh, et al., *arXiv:1106.1706* (2011).
  - [6] R. Laureijs, J. Amiaux, S. Arduini, J. . Auguères, J. Brinchmann, R. Cole, M. Cropper, C. Dabin, L. Duvet, A. Ealet, et al., *arXiv:1110.3193* (2011).
  - [7] R. Ellis, M. Takada, H. Aihara, N. Arimoto, K. Bundy, M. Chiba, J. Cohen, O. Dore, J. E. Greene, J. Gunn, et al., *arXiv:1206.0737* (2012).
  - [8] R. Mandelbaum, A. Slosar, T. Baldauf, U. Seljak, C. M. Hirata, R. Nakajima, R. Reyes, and R. E. Smith, *arXiv:1207.1120* (2012).
  - [9] A. Dekel and O. Lahav, *ApJ* **520**, 24 (1999).
  - [10] T. Matsubara, *Phys.Rev.D* **83**, 083518 (2011).
  - [11] T. Matsubara, *Phys.Rev.D* **77**, 063530 (2008).
  - [12] T. Matsubara, *Phys.Rev.D* **78**, 083519 (2008).
  - [13] P. McDonald, *Phys.Rev.D* **74**, 103512 (2006).
  - [14] T. Okamura, A. Taruya, and T. Matsubara, *JCAP* **8**, 012 (2011).
  - [15] F. Bernardeau, M. Crocce, and R. Scoccimarro, *Phys.Rev.D* **78**, 103521 (2008).
  - [16] T. Matsubara, *Phys.Rev.D*, submitted, *arXiv:1304.????* (2013).
  - [17] T. Matsubara, *Phys.Rev.D* **86**, 063518 (2012).
  - [18] T. T. Nakamura and Y. Suto, *Progress of Theoretical Physics* **97**, 49 (1997).
  - [19] J. P. Henry, *ApJ* **534**, 565 (2000).
  - [20] A. Jenkins, C. S. Frenk, S. D. M. White, J. M. Colberg, S. Cole, A. E. Evrard, H. M. P. Couchman, and N. Yoshida, *MNRAS* **321**, 372 (2001).
  - [21] W. H. Press and P. Schechter, *ApJ* **187**, 425 (1974).
  - [22] J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser, *ApJ* **379**, 440 (1991).
  - [23] R. K. Sheth, H. J. Mo, and G. Tormen, *MNRAS* **323**, 1 (2001).
  - [24] R. K. Sheth and G. Tormen, *MNRAS* **308**, 119 (1999).
  - [25] M. S. Warren, K. Abazajian, D. E. Holz, and L. Teodoro, *ApJ* **646**, 881 (2006).
  - [26] D. S. Reed, R. Bower, C. S. Frenk, A. Jenkins, and T. Theuns, *MNRAS* **374**, 2 (2007).
  - [27] M. Crocce, P. Fosalba, F. J. Castander, and E. Gaztañaga, *MNRAS* **403**, 1353 (2010).
  - [28] M. Manera, R. K. Sheth, and R. Scoccimarro, *MNRAS* **402**, 589 (2010).
  - [29] S. Bhattacharya, K. Heitmann, M. White, Z. Lukić, C. Wagner, and S. Habib, *ApJ* **732**, 122 (2011).

- [30] M. Sato and T. Matsubara, *Phys.Rev.D* **84**, 043501 (2011).
- [31] J. N. Fry and E. Gaztanaga, *ApJ* **413**, 447 (1993).
- [32] T. Baldauf, R. E. Smith, U. Seljak, and R. Mandelbaum, *Phys.Rev.D* **81**, 063531 (2010).
- [33] V. Springel, *MNRAS* **364**, 1105 (2005).
- [34] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. R. Nolte, L. Page, et al., *ApJS* **192**, 18 (2011).
- [35] M. Crocce, S. Pueblas, and R. Scoccimarro, *MNRAS* **373**, 369 (2006).
- [36] P. Valageas and T. Nishimichi, *A&A* **527**, A87 (2011).
- [37] A. Lewis, A. Challinor, and A. Lasenby, *ApJ* **538**, 473 (2000).
- [38] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White, *ApJ* **292**, 371 (1985).
- [39] R. E. Smith, R. Scoccimarro, and R. K. Sheth, *Phys.Rev.D* **75**, 063512 (2007).
- [40] T. Matsubara, *ApJ* **525**, 543 (1999).
- [41] A. Taruya and Y. Suto, *ApJ* **542**, 559 (2000).
- [42] K. Yoshikawa, A. Taruya, Y. P. Jing, and Y. Suto, *ApJ* **558**, 520 (2001).
- [43] U.-L. Pen, *ApJ* **504**, 601 (1998).
- [44] M. Tegmark and B. C. Bromley, *ApJ* **518**, L69 (1999).
- [45] M. Tegmark and P. J. E. Peebles, *ApJ* **500**, L79 (1998).
- [46] U. Seljak and M. S. Warren, *MNRAS* **355**, 129 (2004).
- [47] S. Bonoli and U. L. Pen, *MNRAS* **396**, 1610 (2009).
- [48] J. Carlson, B. Reid, and M. White, *MNRAS* **429**, 1674 (2013).
- [49] J. A. Tyson, F. Valdes, J. F. Jarvis, and A. P. Mills, Jr., *ApJ* **281**, L59 (1984).